

The introduction of symmetry constraints within MaxEnt Jaynes's methodology

F. Holik*

*Center Leo Apostel for Interdisciplinary Studies and, Department of Mathematics,
Brussels Free University Krijgskundestraat 33, 1160 Brussels, Belgium
National University La Plata & CONICET IFLP-CCT, C.C. 727 - 1900 La Plata, Argentina*

C. Massri†

Department of Mathematics, University of Buenos Aires & CONICET IMAS.

A. Plastino‡

*National University La Plata & CONICET IFLP-CCT, C.C. 727 - 1900 La Plata, Argentina
(Dated: today)*

We provide a generalization of the approach to geometric probability advanced by the great mathematician Gian Carlo Rota, in order to apply it to generalized probabilistic physical theories. In particular, we use this generalization to provide an improvement of the Jaynes' MaxEnt method. The improvement consists in providing a framework for the introduction of symmetry constraints. This allows us to include group theory within MaxEnt. Some examples are provided.

Keywords: Maximum Entropy Principle, Geometric Probability, Symmetries in Quantum Mechanics, Generalized Probabilistic Theories

I. INTRODUCTION

Jaynes' MaxEnt approach is a statistical approach in which probability notions become of the essence [1–3]. Thus, new viewpoints regarding probability are susceptible of modifying the MaxEnt approach. We center our present efforts on the notion of geometric probability, characterized by Gian Carlo Rota as the study of invariant measures [4, 5]. This idea has lead to interesting mathematical problems, which have defined a rich field of study. *In this work, we provide a generalization of the Rota's axioms in order to find a physical characterization of the problem of looking for generalized probabilities in the spirit of Jaynes's MaxEnt approach.* As it is well known, this technique relies in the determination of the less unbiased distribution compatible with the known data, by appealing to the maximization of the entropy [1, 2] and has manifold applications in diverse fields of research [6–22] (see [3] for a complete review). Our methodology can be used to find a derivation of both classical and quantum statistical mechanics as well.

Our treatment reformulates the MaxEnt approach in geometric probability terms, allowing for the inclusion of group actions representing physical symmetries. In this framework, states of a physical system are regarded as *invariant measures* over general orthomodular lattices (a lattice is a partially ordered set with unique least upper bounds and greatest lower bounds. For details see, for instance, [23, 24]). The determination of invariant measures under the action of groups representing physical symmetries is of interest in many research fields, as for example, in the problem of the determination of equilibrium states in equilibrium statistical mechanics [25–27].

We also provide an improvement on the treatment of constraints by formulating the problem in the rigorous basis of measure theory, and allowing for them a more general character than mere mean values. We show as well that the introduction of group actions reduces the dimensionality of the mathematical variety on which the maximization process takes place. This economizes computational resources. We demonstrate that this economization can be estimated for certain examples. Finally, we provide some examples and specify conditions under which solutions for our method exist.

The paper is organized as follows. In Section II, we introduce the elementary notions of geometric probability theory following [4][50]. In Section III, we review event structures appearing in both in quantum and classical mechanics —and their associated probabilities—, and in more general probabilistic settings as well. In Section IV, we propose a generalization of geometric probability theory which allows one to describe physical systems. In Section V, we explain how covariance conditions and physical symmetries can be accommodated by our conceptual framework. In Sections VI and VII, we show how to describe, in our framework, quantum coherent states and the correlations appearing in the no-signal polytope. Finally, we draw some conclusions in Section VIII.

II. GEOMETRIC PROBABILITY

In his classical approach to geometric probability [4, 5], Gian Carlo Rota introduces the problem of invariant measures as follows. First, one looks for a measure $\mu : \Sigma \rightarrow \mathbb{R}_{\geq 0}$, defined on a sigma algebra $\Sigma \subseteq \mathcal{P}(\mathbb{R}^n)$, satisfying the following axioms

Axiom 1 (R1)

$$\mu(\emptyset) = 0$$

*Electronic address: holik@fisica.unlp.edu.ar

†Electronic address: cmassri@dm.uba.ar

‡Electronic address: angeloplastino@gmail.com

where \emptyset denotes the empty set. If A and B are measurable sets:

Axiom 2 (R2)

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$$

For Boolean algebras, the above axiom is equivalent to the sum rule

$$\mu(A \cup B) = \mu(A) + \mu(B) \quad (1)$$

for A and B disjoint. The following axiom has to do with the invariance of measures (therefore, the name *invariant measures*):

Axiom 3 (R3) *The measure of a set A does not depend on the position of A ; in other words, if A can be rigidly transformed into B , then, B and A have the same measure.*

Notice that the last axiom involves the action of a group, namely, the Euclidean group E_0 of rotations and translations in Euclidean space. The last axiom specifies a normalization for a given measure; we must pick a special subset and establish its measure. Let us choose the set of parallelotopes P with orthogonal side lengths x_1, \dots, x_n and impose the constrain:

Axiom 4 (R4)

$$\mu(P) = x_1 x_2 \cdots x_n$$

The above axioms yield the usual Lebesgue measure on \mathbb{R}^n . Rota poses the question of what happens if the normalization **Axiom 4** is changed. Instead of **Axiom 4**, one could use one of the following polynomials

$$e_1(x_1, x_2, \dots, x_n) = x_1 + x_2 + \dots + x_n \quad (2a)$$

$$e_2(x_1, x_2, \dots, x_n) = x_1 x_2 + x_1 x_3 + \dots + x_{n-1} x_n \quad (2b)$$

$$\vdots$$

$$e_{n-1}(x_1, x_2, \dots, x_n) = x_2 x_3 \cdots x_n + x_1 x_3 x_4 \cdots x_n + \dots + x_1 x_2 \cdots x_{n-1} \quad (2c)$$

$$e_n(x_1, x_2, \dots, x_n) = x_1 x_2 \cdots x_n \quad (2d)$$

Indeed, the symmetric polynomial e_n is coincident with the normalization of **Axiom 4**. *Geometric probability studies the conditions under which these measures exist, and how they can be used to generate more general ones [4, 5].*

Geometric probability theory can be also used for studying invariant measures in Grassmannians. A complete introduction to the subject can be found in [5]. In the following Section, we will review the formulation of the axioms of a non-commutative probability calculus, i.e., probabilities which generalize Kolmogorov's [28] axioms to non-Boolean settings [29, 30].

III. EVENT STRUCTURES

When faced with a concrete physical problem, we are interested in determining the probabilities of certain events of interest. An event will be the definite outcome of a certain experiment for which we can determine the answer with certainty. As an example, we can think about the detection of a particle (classical or quantal) in a certain region of space-time, and the probability for this event to occur.

It happens that events of a physical system can be endowed with definite mathematical structures [30–32]; if the particle is classical, events may be represented as measurable subsets of the phase space Γ . Measurable subsets of phase space form a well known structure, namely, a *Boolean algebra* [31, 33] that we will denote by $\mathcal{P}(\Gamma)$ [51].

On the other hand, as shown by Birkhoff and von Neumann [34], events associated to a quantum particle will be naturally represented by *projection operators*, specifically those associated to the spectral decomposition of self adjoint operators representing physical observables. Unlike the classical Boolean case, projections of a Hilbert space form an *orthomodular lattice* $\mathcal{P}(\mathcal{H})$, which can be shown to be non-distributive [31, 32, 34] (and thus, not Boolean)[52]. This important mathematical difference between classical and quantum theories is the direct consequence of the incompatibility of complementary observables in QM.

A. Classical case

To illustrate these ideas, let us start by considering the phase space \mathbb{R}^{2n} of a classical system. If f represents an observable quantity, the proposition “the value of f lies in the interval Δ ”, defines an *event* f_Δ , which can be represented as the measurable set $f^{-1}(\Delta)$ (the set of all states which make the proposition true). If the probabilistic state of the system is given by μ , the corresponding probability of occurrence of f_Δ will be given by $\mu(f^{-1}(\Delta))$. As an example, consider the energy of an harmonic oscillator. The proposition “the energy of the oscillator equals ε ” corresponds to an ellipse in phase space for each possible value of ε .

There is a strict correspondence between a classical probabilistic state and the axioms of classical probability theory. Indeed, the axioms of Kolmogorov [28] define a probability function as a measure μ on a sigma-algebra Σ such that

$$\mu : \Sigma \rightarrow [0, 1] \quad (3a)$$

which satisfies

$$\mu(\emptyset) = 0 \quad (3b)$$

$$\mu(A^c) = 1 - \mu(A), \quad (3c)$$

where $(\dots)^c$ means set-theoretical-complement. For any pairwise disjoint denumerable family $\{A_i\}_{i \in I}$,

$$\mu\left(\bigcup_{i \in I} A_i\right) = \sum_i \mu(A_i). \quad (3d)$$

A state of a classical probabilistic theory will be defined as a Kolmogorovian measure with $\Sigma = \mathcal{P}(\Gamma)$. The reader will also notice the analogy between the first two Rota's **Axioms 1 and 2** and the axioms of Kolmogorovian probability theory.

B. Quantum Case

The quantum case can be described in an analogous way. If \mathbf{A} represents the self adjoint operator of an observable associated to a quantum particle, the proposition “the value of \mathbf{A} lies in the interval Δ ” will define an event represented by the projection operator $\mathbf{P}_{\mathbf{A}}(\Delta) \in \mathcal{P}(\mathcal{H})$, i.e., the projection that the spectral measure of \mathbf{A} assigns to the Borel set Δ . The probability assigned to the event $\mathbf{P}_{\mathbf{A}}(\Delta)$, given that the system is prepared in the state ρ , is computed using the Born's rule: $p(\mathbf{P}_{\mathbf{A}}(\Delta)) = \text{tr}(\rho \mathbf{P}_{\mathbf{A}}(\Delta))$. Born's rule defines a measure on $\mathcal{P}(\mathcal{H})$ with which it is possible to compute all probabilities and mean values for all physical observables [31, 34]. As an example, consider the energy of a quantum harmonic oscillator. The proposition “the energy of the oscillator equals ε_i ”, corresponds to the projection operator associated to the eigenspace of the eigenvalue ε_i .

It is well known that, due to Gleason's theorem [35], a quantum state will be defined by a measure s over the orthomodular lattice of projection operators $\mathcal{P}(\mathcal{H})$ as follows [29]:

$$s : \mathcal{P}(\mathcal{H}) \rightarrow [0; 1] \quad (4a)$$

such that:

$$s(\mathbf{0}) = 0 \quad (\mathbf{0} \text{ is the null subspace}). \quad (4b)$$

$$s(P^\perp) = 1 - s(P), \quad (4c)$$

and, for a denumerable and pairwise orthogonal family of projections P_j

$$s\left(\sum_j P_j\right) = \sum_j s(P_j). \quad (4d)$$

C. General Case

Notice that despite their similarities, the difference between (3) and (4) is that Σ is replaced by $\mathcal{P}(\mathcal{H})$, and the other conditions are the natural generalizations of the classical event structure to the non-Boolean setting. A

general probabilistic framework —encompassing the Kolmogorovian and the quantal cases— can be described by the following equations

$$s : \mathcal{L} \rightarrow [0; 1], \quad (5a)$$

(\mathcal{L} standing for the lattice of all events) such that:

$$s(\mathbf{0}) = 0. \quad (5b)$$

$$s(E^\perp) = 1 - s(E), \quad (5c)$$

and, for a denumerable and pairwise orthogonal family of events E_j

$$s\left(\sum_j E_j\right) = \sum_j s(E_j). \quad (5d)$$

where \mathcal{L} is a general orthomodular lattice (with $\mathcal{L} = \Sigma$ and $\mathcal{L} = \mathcal{P}(\mathcal{H})$ for the Kolmogorovian and quantum cases respectively). Eqns. (5a) define what is known as a *generalized probability theory*. Discussing the conditions under which the measure s in Eqns. (5a) is well defined (for very general orthomodular lattices), lies outside the scope of this paper; for a detailed discussion see [32], Chapter 11. It will suffice for us to notice that many examples of interest in physics, including non-relativistic and relativistic QM, and many examples of classical and quantum statistical physics, can be described using orthomodular lattices of projections arising from factors of Type I, II, and III, for which measures such as those defined by Eqs. (5a) are well defined [29, 30].

In the following Sections, we will develop a theoretical framework which combines geometric probability theory, generalized probability theory, and the Jayne's MaxEnt method.

IV. A NEW SET OF AXIOMS FOR PHYSICAL PROBLEMS

A. Classical States As Invariant Measures

Suppose that we are faced with the problem of determining the particular probabilistic state μ of a classical system S . In order to determine μ , we must use the fact that it is a probability measure over the event space $\mathcal{P}(\Gamma)$. Thus, it will obey Eqns. (3), which are equivalent to the first two Axioms of Rotta (Eqs. (1) and (2)) plus the sigma-additivity condition (3d). Imposing **Axiom 3** entails that our system would be in a state which possesses the symmetry of being invariant under the whole group E_0 of translations and rotations of $\mathcal{P}(\Gamma)$. Call E to the group of all possible Galilean transformations acting on the system (notice that $E_0 \subseteq E$). In the general case, the state will not be invariant under all the elements of E_0 , but will be invariant under a subgroup

$G \subseteq E$ (which could be just the identity group, $\{1\}$). For example, equilibrium states of a system with cylindrical symmetry will typically be invariant under rotations and translations along \hat{z} axis, but not for all possible rotations and translations. We will use these observations to generalize Rota's axioms.

Thus, a classical system will have probabilities obeying an alteration of the Rota axioms. In it, i) E_0 in **Axiom (3)** is replaced by a general subgroup $G \subseteq E$, and ii) axiom 4 is replaced by a series of conditions of the form

$$\langle f_i \rangle = r_i, \quad (6)$$

which represent the mean values of observables that are available as empirical data. The group H and conditions (6) represent the a priori information that we have regarding the system (notice that, to the traditional prior information of the Jaynes's method expressed as mean values, we are adding the possibility of symmetry constraints).

Thus, in order to determine the state μ of the system, we must first solve the problem of determining the measures which satisfy the usual probability axioms, plus i) the condition of being invariant under the group G and ii) satisfying the condition given by Eqn. (6). *In this way, the problem of handling geometric probability can be transformed into a physical one.*

B. Quantum States As Invariant Measures

Let us concentrate now on the quantum case before we turn to the general setting. (Continuous) symmetry transformations in QM are represented by the elements of the group of unitary operators \mathcal{U} [36]. If we know in advance that the state that we are looking for possesses a certain symmetry, this condition will be represented by the invariance of the state under the action of a subgroup $\mathcal{G} \subseteq \mathcal{U}$. Next, a series of conditions on mean values of observables can be added. These can be either mean values of operators or more general ones, but which are insufficient on their own to fully determine the state. These conditions can be cast in the form

$$\langle \mathbf{A}_i \rangle = a_i \quad (7)$$

A state will be represented by a *measure* s over the event structure $\mathcal{P}(\mathcal{H})$. In other words, we are looking for a measure s which i) satisfies Eqns. (4), ii) that is *invariant* under the action of the group \mathcal{G} , and iii) satisfies Eqns. (7). Thus, in order to determine a quantum state compatible with the prior knowledge about symmetries and mean values, we must determine a measure such that the **Axioms (3)** and (4) be adequately modified.

We see that, as in the classical case, the Rota's problem can be extended to the problem of determining the state of a physical system, provided we generalize subsets of Euclidean space to the lattice of projections in a Hilbert space, replace the roto-translational group by the

corresponding quantum one, and replace the normalization condition by known mean values of a given set of observables. These conditions restrict the possible states to a subset of the space of quantum states. Following Jaynes [2] now, the least biased probability distribution can be determined by maximizing von Neumann's entropy in this subset. It is nice that these observations are susceptible of an even greater degree of generalization.

C. Invariant Measures In Generalized Theories

Now we pass to a systematic generalization of the above procedure for quite arbitrary statistical theories, which will provide a new ground for the MaxEnt principle. In this vein, we are led to formulate the following set of axioms for a general physical system, incorporating prior knowledge about symmetries and conditions on expectation values (or even more general conditions). The objective is to determine the unknown state s of given system as an invariant measure obeying Eqs. (4).

Symmetries: Knowledge about symmetries of the physical system will be represented by the existence of a subgroup \mathfrak{G} of the group automorphisms of \mathcal{L} , $Aut(\mathcal{L})$, such that for all $g \in \mathfrak{G}$, and for all $E \in \mathcal{L}$,

$$s(g \cdot E) = s(E). \quad (8)$$

Normalization condition: There exists a set of equations $\{e_i\}_I$ in the values $\{s(E_j)\}_J$,

$$e_i(s(E_1), s(E_2), \dots) = 0, \quad (9)$$

where $\{E_j\}_J \subseteq \mathcal{L}$ is some subset of events.

To summarize, we set down all the axioms that the unknown state ν —now considered as a generalized invariant measure $\nu : \mathcal{L} \rightarrow [0; 1]$ over an arbitrary orthomodular lattice \mathcal{L} —must satisfy:

Axiom 5 (G1)

$$\nu(\emptyset) = 0$$

Axiom 6 (G2)

$$\nu(E^\perp) = 1 - s(E)$$

Axiom 7 (G3) For a denumerable and pairwise orthogonal family of events E_j ,

$$\nu\left(\sum_j E_j\right) = \sum_j \nu(E_j)$$

Axiom 8 (G4) For all $g \in \mathfrak{G}$

$$\nu(g \cdot E) = \nu(E)$$

Axiom 9 (G5) *There exists a family of events $\{E_j\}$ which satisfy the equations defined by functions e_i*

$$e_i(\nu(E_1), \nu(E_2), \dots, \nu(E_{m_i})) = 0$$

48625636 The above **Axioms** represent our generalization of geometric probability to the noncommutative case. **Axioms** (5), and (6) and (7) univocally determine a convex set \mathcal{S} (provided that ν be well defined, cf. [32], Chapter 11). It is important to remark that the introduction of **Axiom** (8) yields a smaller set $\mathcal{S}_{\mathfrak{G}} \subseteq \mathcal{S}$ which is also convex. The addition of **Axiom** (9) determines a manifold \mathcal{M} , which, when intersected with $\mathcal{S}_{\mathfrak{G}}$, will not necessarily yield a convex set. However, it can be shown that if the constraints are mean values imposed on observables, or more generally, on effects, the set determined by $\mathcal{S}_{\mathfrak{G}} \cap \mathcal{M}$ will be convex [37]. Thus, the set of states compatible with the prior knowledge about symmetries and measured quantities will be the intersection $\mathcal{S}_{\mathfrak{G}} \cap \mathcal{M}$.

Once this set is determined, Jaynes's entropic maximization process singles out the less unbiased state which will rule the probabilities of the system. In the following **Section**, we discuss which entropic measures are to be used for this purpose. Notice that if \mathcal{S} is compact, then $\mathcal{S}_{\mathfrak{G}}$ and $\mathcal{S}_{\mathfrak{G}} \cap \mathcal{M}$ will be also compact, and we can ensure the existence of a solution for the maximization procedure (provided that the entropic measure that we use be continuous). Many physical examples comply with these assumptions (for example, in non-relativistic quantum mechanics, the state space is compact and the symmetry groups are locally compact).

D. Entropies

We wish to define a meaningful notion of entropy for using it in several frameworks, in the sense of being applicable to QM, classical mechanics, and to general theories. Thus, we need an appropriate notion of information measure, to be applied to general statistical theories. One possibility is to use the so called *measurement entropy*, which reduces to Shannon's measure for classical models and to von Neumann's in the quantum case [38, 39]. Let s be a state in a generalized probability theory. Then, following Ref. [38], we define

$$H_E(\nu) := - \sum_{x \in E} \nu(x) \ln(\nu(x)), \quad (10)$$

$$H(\nu) := \inf_{E \in \mathcal{L}} H_E(\nu). \quad (11)$$

We show a comparison of the different cases in Table I.

E. Frame Functions And Group Actions

Assume that a group \mathfrak{G} is acting by automorphisms on a lattice of events \mathcal{L} , $G \subseteq \text{Aut}(\mathcal{L})$ [36, 40]. Consider the

	CLASSICAL	QUANTUM	GENERAL
LATTICE	$\mathcal{P}(\Gamma)$	$\mathcal{P}(\mathcal{H})$	\mathcal{L}
GROUP	$G \subseteq E$	$\mathcal{G} \subseteq \mathcal{U}$	$\mathfrak{G} \subseteq \text{Aut}(\mathcal{L})$
ENTROPY	$-\sum_i p(i) \ln(p(i))$	$-\text{tr} \rho \ln(\rho)$	$\inf_{E \in \mathcal{L}} H_E(\nu)$

TABLE I: Table comparing the differences between the classical, quantal, and general cases.

convex set \mathcal{S} of Section IV C. **Axiom** (8) states that invariant states are constant along the orbits of the action,

$$s(g \cdot E) = s(E), \quad g \in \mathfrak{G}, E \in \mathcal{L},$$

and an invariant state in \mathcal{L} defines in a canonical way a state in \mathcal{L}/\mathfrak{G} , where \mathcal{L}/\mathfrak{G} is the quotient lattice.

Assume now that the lattice \mathcal{L} is atomic, where the set of atoms is an n -dimensional compact manifold \mathcal{A} . According to Gleason [35], a state in \mathcal{L} is determined by a *frame function* in \mathcal{A} , that is,

$$f : \mathcal{A} \rightarrow \mathbb{R}, \quad \sum_{i=1}^r f(x_i) = 1,$$

where $\{x_1, \dots, x_r\}$ is a set in \mathcal{L} such that $x_i \perp x_j$ ($i \neq j$) and $x_1 \vee \dots \vee x_r = 1$. Call \mathcal{F} to the set of frame functions. The full group of automorphisms of the atomic lattice, $\text{Aut}(\mathcal{L})$, induces an action in \mathcal{A} and \mathcal{F} is stable under this action. If $f \in \mathcal{F}$ and $g \in \text{Aut}(\mathcal{L})$, then $g \cdot f$ is also a frame function. Note that the continuous frame functions $\mathcal{F}_{\text{cont}} \subseteq \mathcal{F}$ is a subset of all the bounded continuous functions in \mathcal{A} , $\mathcal{F}_{\text{cont}} \subseteq L^\infty(\mathcal{A})$, and that the polynomial frame functions are dense in $\mathcal{F}_{\text{cont}}$.

The action of the group G in \mathcal{L} restricts itself to an action on \mathcal{A} , and a frame function determines an invariant state if and only if the frame function is invariant, $g \cdot f = f$, for all $g \in \mathfrak{G}$. Thus, the invariant states are characterized by the frame functions in \mathcal{A}/\mathfrak{G} . Recall that the dimension of \mathcal{A}/\mathfrak{G} is equal to the dimension of \mathcal{A} minus the dimension of an orbit.

As an example, consider an $(n+1)$ -dimensional Hilbert space and its lattice of subspaces, \mathcal{L} . The set of atoms (the rays in the Hilbert space) is a projective space \mathbb{P}^n . It is a compact variety of dimension n , $\mathcal{A} = \mathbb{P}^n$. The full group of automorphisms of \mathcal{L} is the Lie group \mathcal{U} . In [35], the fact that the set of frame functions \mathcal{F} is stable under \mathcal{U} is used to characterize frame functions as density matrices (positive semi-definite self-adjoint operators of the trace class).

Consider now a group $\mathcal{G} \subseteq \mathcal{U}$, acting on \mathbb{P}^n , and let us consider states invariant under the group \mathcal{G} . Given that states are characterized by density matrices, the invariant states are density matrices stable under \mathcal{G}

$$\rho = g \cdot \rho, \quad \forall g \in \mathcal{G},$$

or equivalently, frame functions in \mathbb{P}^n/\mathcal{G} . Note that we are reducing the dimension of the convex set of states and the reduction will depend on the nature of the action of \mathcal{G} .

V. COVARIANCE AND SYMMETRIES

A space time symmetry will have an action on the observables of the system and on the state space. But this implies at the same time that it will have an action on the associated operational logic. As an example, consider the Galilei group in non-relativistic QM. Any operator of the group acts on the variety of space time observables (position, momentum) but at the same time there exists a representation of this group in the set of unitary operators of Hilbert space. Indeed, the content of Wigner's theorem asserts that symmetry transformation preserving probabilities will have a representation as a unitary or anti-unitary operator in Hilbert space. This means that for each symmetry, say, a rotation, there exists an automorphism acting on the logic of projection operators.

Thus, symmetries are usually generalized as follows [53]. Suppose that we have a group \mathfrak{G} representing symmetries of a physical system. Call \mathcal{S} the set of all probability measures. The elements of \mathfrak{G} will also induce transformations in \mathcal{S} as convex automorphisms. As it is well known [36, 40], this group will also have a representation in $\text{Aut}(\mathcal{L})$. Thus, for any element $g \in \mathfrak{G}$, any event $E \in \mathcal{L}$ and any $\nu \in \mathcal{S}$, a symmetry of the system will satisfy the covariance condition

$$\nu(E) = \nu'(E'), \quad (12)$$

where $E' = g \cdot E$ and $\nu' = g \cdot \nu$.

The above equation is important for two main reasons:

- It allows us to incorporate into our system the very important notion of representation of groups, acting as convex automorphisms on \mathcal{S} and automorphisms of \mathcal{L} . The action of these groups represents the actions of symmetry transformations (including the spatiotemporal ones) and imposes conditions on the geometry of \mathcal{S} and observable algebras.
- We will use this approach to define coherent states in the general setting. First, because the introduction of symmetries obeying the covariance condition (12) allows for the definition of a base state (as is the case for the vacuum state of the electromagnetic field). Secondly, because the group axiom allows us to pick up only those measures which satisfy the condition of being coherent states.

VI. COHERENT STATES

Given the Heisenberg uncertainty relation in a state ρ

$$\Delta \mathbf{P} \Delta \mathbf{Q} \geq \frac{\hbar}{2} \quad (13)$$

where for an operator \mathbf{O} , $\Delta \mathbf{O} = \sqrt{\langle \mathbf{O}^2 \rangle - \langle \mathbf{O} \rangle^2}$, coherent states [41–43] are defined as those which saturate (13) with equal mean values, i.e.:

$$\Delta \mathbf{P} \Delta \mathbf{Q} = \frac{\hbar}{2} \quad (14a)$$

$$\Delta \mathbf{P} = \sqrt{\frac{\hbar}{2}} = \Delta \mathbf{Q} \quad (14b)$$

Thus, we can easily incorporate such states into our conceptual framework by replacing (8) by Eqs. (14). Note that Eq. (14b) produces a real algebraic variety \mathcal{M} in the real vector space of Hermitian operators (it is given by the zero locus of the two polynomial equations of degree two, $\Delta \mathbf{P} = \sqrt{\frac{\hbar}{2}}$ and $\Delta \mathbf{Q} = \sqrt{\frac{\hbar}{2}}$).

In arbitrary dimension ($n \leq \infty$) the states satisfying Eq. (14b) are given by the intersection of two quadrics. Recall that any quadric can be parameterized and thus the intersection $\mathcal{C} \cap \mathcal{M}$ can be computed in finite dimensions. If the convex set \mathcal{C} is a compact set, then the intersection $\mathcal{C} \cap \mathcal{M}$ is also compact.

We can also define coherent states using group theory. This has the advantage of being easily applicable to general statistical theories [54]. While the choice of a reference state s_0 is, in principle, arbitrary [43], the use of physical symmetries could be useful for its determination. These will be represented by a group action \mathfrak{G} which, as mentioned above, induces actions in \mathcal{L} and \mathcal{S} . This procedure singles out the correct reference state s_0 [43] by using the generalization of geometric probability described in previous Sections. Once s_0 is specified, we invoke the action of a given dynamical group G , determine its maximum stability subgroup H [43], and construct the set of all coherent states $\mathcal{S}_G \subseteq \mathcal{S}$ as follows

$$s_g := g \cdot s_0, \quad (15)$$

where g ranges over all the elements of G/H [43].

VII. BELL INEQUALITIES, NO-SIGNAL POLYTOPE AND LOCAL POLYTOPE

Immense interest generates in the study of correlations in QM. For two separate observers, A and B , both of them having available two observables $\{a_0, a_1\}$ and $\{b_0, b_1\}$, with two possible outcomes for each, the correlations will be governed by probability distributions of the form $P(a_i, b_j | x, y)$. It can be shown that the following inequalities can be violated by QM

$$S = |\langle a_0 b_0 \rangle + \langle a_1 b_0 \rangle + \langle a_0 b_1 \rangle - \langle a_1 b_1 \rangle| \leq 2, \quad (16)$$

These are known as the Clauser-Horne-Shimony-Holt (CHSH) inequalities [44, 45]. The no-signal polytope, formed by all possible correlations respecting the no-signal condition of special relativity, is defined by the following conditions [45]

$$\sum_j P(a_i, b_j | x, y) = \sum_j P(a_i, b_j | x, y') \quad \forall y, y' \quad (17a)$$

$$\sum_i P(a_i, b_j | x, y) = \sum_i P(a_i, b_j | x', y) \quad \forall y, y' \quad (17b)$$

Quantum correlations can violate the CSHS inequalities, but at the same time, they respect the no-signal condition (the distributions $P(a, b | x, y)$ lie inside the no-signal polytope). One may ask which is the characteristic trait of quantum mechanics that distinguishes it from general statistical theories which are also no-signal, but do not produce the correlations predicted by QM [45]. This issue can be studied within our theoretical framework by setting conditions (16) and (17) as axioms in the event space. By replacing condition (9) by (16), we obtain the local polytope, and by replacing it by (17), we obtain the no-signal polytope. The reformulation of these geometrical objects within our framework could permit the study of the action of suitable groups of space-time symmetries (by introducing these groups through Axiom (8)).

VIII. CONCLUSIONS

It is important to remark that a systematic presentation of the Jaynes's method, as we have done here, has not yet been advanced in the literature, as far as we know. We summarize our conclusions as follows:

- The use of a formulation based on the traditional axiomatic method of measure theory allows for a rigorous approach to MaxEnt.
- We have shown that many cases fall within the axiomatic framework presented here (coherent states, no-signal polytopes, local polytopes). When our group symmetry is reduced to the identity and the constraints are expressed as mean values, our method reduces to previous generalizations of the Jaynes's methodology [37, 39].
- When $\mathcal{L} = \mathcal{L}_{v\mathcal{N}}$ or $\mathcal{L} = \mathcal{P}(\Gamma)$, and the constraints are expressed as mean values, our method reduces to the pioneer Jaynes's one for the quantum and classical cases, respectively.

- Our rigorous formulation allows us to establish precise conditions for the existence of solutions to the MaxEnt problem for very general constraints (including group theory, non-linear conditions on the mean values of observables, and inequalities as well).
- At the same time, we provide an intrinsic geometric characterization for the different mathematical objects defined within our theoretical framework (quadratics for coherent states, a convex set for the local polytope, etc.). Notice that this may be of help in studying the geometrical properties of the non-signal and local polytopes for the most general case (a continuous range of observables with possibly continuous spectra. In QM, in infinite dimensional Hilbert spaces). Our formulation may help to extrapolate, in the future, results from Geometric Probability Theory to physics.
- By reformulating the problem in terms of the determination of invariant measures, we provide a natural framework for the introduction of group theory. We have explicitly shown that the introduction of groups reduces the dimensionality of the mathematical variety in which the maximization process takes place. Thus, our proposal may be useful to economize computational resources. Our axioms allow one to incorporate into the Jaynes's framework the symmetries of the physical system under study. For example, one could insert a group representing a spacial symmetry of a system. This method yields a powerful resource for deriving laws of physics out of general physical principles.
- The facts that i) probability theory is a well established theory and ii) explicit solutions to our problem can be found (as in the examples studied in this work) show that our mathematical problem is meaningful. This fact constitutes a clear improvement on the MaxEnt method, giving a step forward into its axiomatization.

Acknowledgments

This work has been supported by CONICET.

-
- [1] E. Jaynes, Phys. Rev. **106**, 620 (1957).
 - [2] E. Jaynes, Phys. Rev. **108**, 171 (1957).
 - [3] S. Pressé, K. Ghosh, J. Lee, and K. Dill, Rev. Mod. Phys. **85**, 1115 (2013).
 - [4] G. C. Rota, The Mathematical Intelligencer **20**, 11 (1998).
 - [5] D. A. Klain and G. C. Rota, *Introduction to geometric probability*, Lezioni Lincee. [Lincei Lectures] (Cambridge University Press, Cambridge, 1997).
 - [6] A. Cavagna, I. Giardina, F. Ginelli, T. Mora, D. Piovani, R. Tavarone, and A. M. Walczak, Phys. Rev. E **89**, 042707 (2014).
 - [7] G. P. Beretta, Phys. Rev. E **90**, 042113 (2014).
 - [8] R. Sinatra, J. Gómez-Gardeñes, R. Lambiotte, V. Nicosia, and V. Latora, Phys. Rev. E **83**, 030103 (2011).
 - [9] A. Dirks, P. Werner, M. Jarrell, and T. Pruschke, Phys. Rev. E **82**, 026701 (2010).
 - [10] M. Trovato and L. Reggiani,

- Phys. Rev. E **81**, 021119 (2010).
- [11] L. Diambra and A. Plastino, Phys. Rev. E **52**, 4557 (1995).
- [12] L. Rebollo-Neira and A. Plastino, Phys. Rev. E **65**, 011113 (2001).
- [13] L. Diambra and A. Plastino, Phys. Rev. E **53**, 1021 (1996).
- [14] G. Goswami and J. Prasad, Phys. Rev. D **88**, 023522 (2013).
- [15] R. Cofré and B. Cessac, Phys. Rev. E **89**, 052117 (2014).
- [16] N. Lanatà, H. U. R. Strand, Y. Yao, and G. Kotliar, Phys. Rev. Lett. **113**, 036402 (2014).
- [17] M. Trovato and L. Reggiani, Phys. Rev. Lett. **110**, 020404 (2013).
- [18] D. Gonçalves, C. Lavor, M. Gomes-Ruggiero, A. Cesário, R. Vianna, and T. Maciel, Phys. Rev. A **87**, 052140 (2013).
- [19] N. Canosa, A. Plastino, and R. Rossignoli, Phys. Rev. A **40**, 519 (1989).
- [20] L. Arrachea, N. Canosa, A. Plastino, M. Portesi, and R. Rossignoli, Phys. Rev. A **45**, 7104 (1992).
- [21] G. Tkačik, O. Marre, T. Mora, D. Amodei, M. J. BerryII, and W. Bialek, Journal of Statistical Mechanics: Theory and Experiment **2018**, 03011K (2018).
- [22] K. H. Knuth, (2014), arXiv:1411.1854 [quant-ph].
- [23] G. Kalmbach, *Orthomodular lattices*, London Mathematical Society Monographs, Vol. 18 (Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, 1983).
- [24] F. Holik, C. Massri, A. Plastino, and L. Zuberaman, International Journal of Theoretical Physics **52**, 1836 (2013).
- [25] O. E. Lanford and D. W. Robinson, Journal of Mathematical Physics **9**, 1120 (1968).
- [26] I. Lanford, Oscar E. and D. Robinson, Communications in Mathematical Physics **9**, 327 (1968).
- [27] O. E. Lanford and D. Ruelle, Comm. Math. Phys. **13**, 194 (1969).
- [28] A. N. Kolmogorov, in *Probability theory and mathematical statistics (Tbilisi, 1982)*, Lecture Notes in Math., Vol. 1021 (Springer, Berlin, 1983) pp. 1–5.
- [29] M. Rédei and S. J. Summers, Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics **38**, 390 (2007).
- [30] M. Rédei, *Quantum logic in algebraic approach*, Fundamental Theories of Physics, Vol. 91 (Kluwer Academic Publishers Group, Dordrecht, 1998).
- [31] M. Dalla Chiara, R. Giuntini, and R. Greechie, *Reasoning in quantum theory*, Trends in Logic—Studia Logica Library, Vol. 22 (Kluwer Academic Publishers, Dordrecht, 2004) sharp and unsharp quantum logics.
- [32] E. G. Beltrametti and G. Cassinelli, *The logic of quantum mechanics*, Encyclopedia of Mathematics and its Applications, Vol. 15 (Addison-Wesley Publishing Co., Reading, Mass., 1981) with a foreword by Peter A. Carruthers.
- [33] G. Boole, *An investigation of the laws of thought*, Cambridge Library Collection (Cambridge University Press, Cambridge, 2009) on which are founded the mathematical theories of logic and probabilities, Reprint of the 1854 original, Previously published by Dover Publications, Inc., New York, 1957; Prometheus Books, Amherst, NY, 2003.
- [34] G. Birkhoff and J. von Neumann, Ann. of Math. (2) **37**, 823 (1936).
- [35] A. Gleason, in *The Logico-Algebraic Approach to Quantum Mechanics*, The University of Western Ontario Series in Philosophy of Science, Vol. 5a, edited by C. Hooker (Springer Netherlands, 1975) pp. 123–133.
- [36] V. S. Varadarajan, *Geometry of quantum theory. Vol. I* (D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London, 1968) the University Series in Higher Mathematics.
- [37] F. Holik and A. Plastino, Journal of Mathematical Physics **53**, 073301 (2012).
- [38] H. Barnum, J. Barrett, L. O. Clark, M. Leifer, R. Spekkens, N. Stepanik, A. Wilce, and R. Wilke, New Journal of Physics **12**, 033024 (2010).
- [39] C. Hein, Foundations of Physics **9**, 751 (1979).
- [40] V. S. Varadarajan, *Geometry of quantum theory. Vol. II* (Van Nostrand Reinhold Co., New York-Toronto, Ont.-London, 1970) the University Series in Higher Mathematics.
- [41] R. Glauber, Phys. Rev. **130**, 2529 (1963).
- [42] R. Glauber, Phys. Rev. **131**, 2766 (1963).
- [43] W.-M. Zhang, D. H. Feng, and R. Gilmore, Rev. Mod. Phys. **62**, 867 (1990).
- [44] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
- [45] S. Popescu, Nature **10**, 264 (2014).
- [46] R. J. von Steudt, and J. Skilling, Axioms **1**, 38 (2012).
- [47] G. Ludwig, *Foundations of quantum mechanics. I*, Texts and Monographs in Physics (Springer-Verlag, New York, 1983).
- [48] G. Ludwig, *Foundations of quantum mechanics. II*, Texts and Monographs in Physics (Springer-Verlag, New York, 1985).
- [49] G. W. Mackey, *Mathematical foundations of quantum mechanics* (Dover Publications, Inc., Mineola, NY, 2004) with a foreword by A. S. Wightman, Reprint of the 1963 original.
- [50] The reader familiarized with this theory can skip this Section.
- [51] A Boolean lattice will be a partially ordered set for which i) the least upper bound (disjunction) and maximum lower bound (conjunction) exists for every pair of elements; ii) it is orthocomplemented; iii) it is distributive. A typical model for a Boolean lattice will be that of the subsets of a given set, with set intersection as conjunction, set union as disjunction, and set theoretical complement as orthocomplementation [31] (see also [46] for a study of the algebraic symmetries of Boolean lattices).
- [52] An orthomodular lattice will be an orthocomplemented lattice for which a condition weaker than distributivity holds (see for example [23, 24, 29]). Boolean algebras are always orthomodular lattices, but the converse is not true [31]. For $\mathcal{P}(\mathcal{H})$, conjunction is given intersection, disjunction by closure of direct sum, and orthocomplementation by orthogonal complement of the closed subspaces associated to each projection operator (projection operators can be put in one to one correspondence with closed subspaces of \mathcal{H}).
- [53] This methodology can be traced back to [47], [48], [36], [40] and [49].
- [54] It is important to remark here that both the definition of coherent states that uses Eqs. (14) and the group theoretical one are equivalent for the case of the electromagnetic field, but will not be equivalent in general (as is the case for finite dimensional Hilbert spaces) [43]. Thus, it is not expected that these definitions will be equivalent in arbitrary statistical theories neither.